Bijlage 2 Normeringssystematiek centrale examens vo

Inleiding
Op verzoek van de staatssecretaris van OCW zet het College voor Examens (CvE) hierbij uiteen hoe het zijn wettelijke taak uitvoert om de normen voor de omzetting van scores in cijfers bij de centrale examens (CE) te bepalen. Bij het normeren van de CE's maakt het CvE gebruik van een normeringsterm. De normeringsterm (N-term) bepaalt welk cijfer bij een score hoort. De N-term corrigeert de onvermijdelijke verschillen in moeilijkheidsgraad tussen examens van opeenvolgende jaren, zodat de prestatie-eis (het resultaat dat nodig is voor een voldoende) van jaar tot jaar gelijk blijft.
Dus: door te draaien aan de knop van de N-term (equivaleren in vaktermen) bereikt het CvE dat aan kandidaten in 2013 dezelfde eisen worden gesteld voor het behalen van een voldoende als in de jaren daarvoor.

Vaardigheidsverschillen worden zichtbaar
De normeringsterm wordt *niet* gebruikt om het aantal of percentage geslaagde kandidaten te beïnvloeden. Dat is echter *wel* het meest verbreide misverstand. [1] Hoe komt dat? Het niveau van de Nederlandse schooljeugd is doorgaans relatief stabiel. Dat verklaart dat de N-termen over meerdere jaren aanleiding geven tot vergelijkbare percentages voldoende en gemiddelde cijfers. Dat zegt mogelijk iets over de continuïteit van onderwijskwaliteit. Het kan natuurlijk ook voorkomen dat het vaardigheidsniveau stijgt of daalt. Zo is recent zichtbaar geworden dat de *gemiddeld-voldoende-eis* die in 2012 voor het CE werd ingevoerd en de *kernvakkenregel* in 2013, hun doel niet hebben gemist. Uit de equivaleringstechnieken die Cito in opdracht van het CvE heeft verricht bleken de populaties examenkandidaten van 2012 en 2013 vaardiger dan die in jaren vóór de aanscherping van de uitslagregels.

Normeringsproces in een notendop
Hoe is het CvE in staat dit te meten?
Daarvoor is de rolverdeling tijdens het normeringsproces van belang.
- de kandidaten maken hun examen; de correctoren stellen de scores vast en geven die door aan Cito (WOLF-score).
- Cito analyseert, rekent en vergelijkt. Die opbrengst wordt getoetst aan resultaten van aanvullend onderzoek door middel van pretests en posttests. Dit totaal levert de brandstof voor een technisch normeringsadvies aan het CvE.
- CvE vraagt aanvullend advies aan elke vaststellingscommissie, die het examen vaststelde. Bovendien checkt CvE of er via de Examenlijn of van LAKS of nog bijzonderheden zijn waar bij de vaststelling van de N-term rekening mee gehouden moet worden. Algemene stelregel daarbij: de kandidaten mogen niet het dupe worden van onvolkomenheden.
- De CvE-leiding (voorzitter, directeur en sectormanagers) stelt de N-term per vak en schooltype vast.

Rekentoets VO
De rekentoets in het VO (en de centrale examens taal en rekenen MBO) wordt genormeerd aan de hand van een vergelijkbaar proces. Maar door de aanvullende eis –onderlinge vergelijkbaarheid van bijvoorbeeld niveau 3F in havo/vwo en MBO-niveau-4 - is er voor gekozen met behulp van een vaardigheidsschaal een tussenstap te maken in de omzetting van score naar cijfer.

Verbindt deze vaardigheidsschaal het niveau door de jaren net als de N-term?
Ja. Maar er is nog wat extra tijd nodig. Simpelweg omdat er bij de rekentoets nog weinig verleden is opgebouwd. Wat helpt, is dat recent referentiesets konden worden vastgesteld en genormeerd. Deze vormen als het ware de verankering van het niveau en zorgen voor een gelijkwaardige

---

[1] Zou er bij de normering worden gemikt op een vast gemiddelde dat door de jaren heen gelijk blijft, dan zou een daling of stijging van het vaardigheidsniveau nooit zichtbaar worden in de examenuitslag. De normering zou dan immers met de vaardigheid van de kandidaten mee dalen of stijgen.

normering van de centrale toetsen/examens over de referentieniveaus rekenen in de verschillende sectoren.

Het CvE is graag bereid om ten behoeve van de Tweede Kamer een briefing over de normeringssystematiek te verzorgen.

Toelichting – De normering van de centrale examens VO

Bij de normering van de centrale examens VO worden drie stappen doorlopen;

- een technisch normeringsadvies van Cito;
- een advies van iedere CvE-vaststellingscommissie;
- vaststelling door voorzitter, directeur en sectormanagers van het College voor Examens (CvE).

Stap 1: het technisch normeringsadvies

Na de afname maakt Cito in opdracht van het CvE van ieder centraal examen een toets- en itemanalyse. Dit gebeurt op basis van de scores die de correctoren per vraag aan de kandidaten hebben toegekend. In een toets- en itemanalyse is zichtbaar hoeveel scorepunten de kandidaten gemiddeld hebben behaald, zowel per vraag als voor het gehele examen. Op het niveau van de toets blijkt daarnaast onder meer hoe de scorepunten gespreid zijn en hoe betrouwbaar het examen als meting was. De toets- en itemanalyse is daarmee de kwantitatieve graadmeter voor hoe het examen 'gevallen' is. Cito relateert de gegevens uit de toets- en itemanalyse aan het referentie-examen (zie kader) en – indien voorhanden – aan de gegevens uit de pretest (zie kader) of posttest (zie kader) over het examen. Daarmee komt Cito tot zijn technisch normeringsadvies. Dit technisch normeringsadvies heeft de vorm van de voor het scholenveld bekende normeringsterm of N-term (zie kader).

Stap 2: het advies van de CvE-vaksectie

Op de dag voordat de N-termen aan de scholen bekend gemaakt worden komen alle vaststellingscommissies van het CvE in het Cito-gebouw bijeen voor de normeringsvergadering. Iedere vaststellingscommissie beschikt dan over de toets- en itemanalyse en het technisch normeringsadvies van Cito en wordt bijgestaan door een toetsdeskundig vakmedewerker van Cito. De vaststellingscommissie beschikt bovendien over de – kwalitatieve – reacties op het examen die zijn binnengekomen via de CvE-examenlijn en over de verslagen van examenbesprekingen die door vakverenigingen georganiseerd zijn. De CvE-vaststellingscommissie weegt alle reacties en brengt een advies N-term uit. Stel bij voorbeeld dat vakdocenten aangeven dat veel leerlingen in tijdnood gekomen zijn en de toets- en itemanalyse bevestigt dat, dan kan de vaststellingscommissie dit in haar advies over de N-term verdisconteren. Een ander voorbeeld: als uit de toets- en itemanalyse zou blijken dat een bepaalde opgave door de minst vaardige kandidaten juist goed en door de vaardige kandidaten slecht is gemaakt, zal de vaksectie die vraag nog eens extra onder de loep nemen. Achteraf is dan meestal wel duidelijk hoe deze vraag beter gesteld had kunnen worden. Bij haar advies over de N-term houdt de vaksectie rekening met eventuele minder valide vragen en voor toekomstige examens wordt daar lering uit getrokken. Voor de vaksecties en de toetsdeskundigen van Cito is de normeringsvergadering daardoor een belangrijk evaluatiemoment.

Stap 3: de vaststelling van de N-termen

Voorzitter, directeur en sectormanagers stellen op basis van deze stappen de N-termen vast. Op hun beurt worden zij daarin bijgestaan door de sectormanagers en psychometrici van Cito. Sectormanagers van CvE en Cito hebben bovendien vooraf nog met het LAKS gesproken, zodat behalve van de vakverenigingen, ook de bevindingen van de leerlingenorganisatie bekend zijn. In juni 2010 zijn zo voor de centrale examens VO van het eerste tijdvak 224 N-termen vastgesteld: 95 beroepsgerichte vakken vmbo, 55 algemene vakken vmbo, 32 havo, 42 vwo. Daarnaast zijn N-termen vastgesteld voor de flexibele en digitale centrale examens in de basis- en kaderberoepsgerichte leerweg.

Op de dag na de normeringsvergadering worden alle N-termen 's ochtends om 8 uur bekend gemaakt via Examenblad.nl. Dit stelt de VO-scholen in staat om van hun kandidaten de cijfers voor de centrale examens en de uitslag van het eindexamen vast te stellen.
Later diezelfde dag verschijnen dan de eerste vlaggen met schooltassen.

Een *referentie-examen* is een centraal examen uit een eerder examenjaar dat alom gezien wordt als een goede operationalisering van de prestatie-eisen die voor het desbetreffende vak gelden.

In een *pretest* worden combinaties van opgaven van een toekomstig centraal examen afgenomen en een referentie-examen afgenomen. Cito kan daardoor de vaardigheid van de pretest-kandidaten vergelijken met de vaardigheid van twee groepen kandidaten: de groep kandidaten die het referentie-examen aflegden en, na de afname, de populatie die het gepresteste centraal examen heeft gemaakt. Door jaarlijks te pretesten ontstaat een cijfermatig beeld van eventuele veranderingen in de vaardigheid van examenkandidaten over de jaren heen.

Een *posttest* werkt net als een pretest, maar vindt kort *na* het centraal examen plaats.

Pretest en posttest zijn methoden om examens te *equivaleren.* Met de inzet van equivaleringsmethoden wil het CvE zeker stellen dat over de jaren heen gelijke prestatie-eisen aan kandidaten kunnen worden gesteld.

Een *normeringsterm* of *N-term* is een getal dat voor alle mogelijke scores voor een centraal examen bepaalt welk cijfer daarbij hoort.
Bij N = 1,0 moet de leerling moet 50% van de maximale score behalen voor een voldoende cijfer. Bij N = 0,0 is dat ongeveer 60% en bij N = 2,0 ongeveer 40%.

Een toelichting op deze normeringssystematiek staat op www.cve.nl: kies 'Onderwerpen', vervolgens 'Centrale examens VO' en daarna 'Examencyclus'. Klik tenslotte op 'Toetsspecial normering'.

Toelichting - De normering van de Rekentoets VO

De rekentoets VO wordt digitaal afgenomen. Tijdens een door het CvE aangewezen afnameperiode van twee weken bepaalt de school zelf de afnametijdstippen. Daartoe krijgen alle scholen de beschikking over tien varianten van de rekentoets.
Als gevolg van de flexibele en digitale afname wijkt de omzetting van scores in cijfers bij de rekentoets VO af van de wijze van normeren bij de centrale examens in het voortgezet onderwijs. Er wordt *niet* gewerkt met normeringstermen (N-termen), maar met een vaardigheidsschaal. *Per variant* van de rekentoets VO wordt iedere score gekoppeld aan een vaardigheidsniveau. Vervolgens is aan elk vaardigheidsniveau een cijfer gekoppeld, met inachtneming van de cesuur. Door de omzetting van scores in cijfers te laten verlopen via een vaardigheidsschaal, wordt ervoor gezorgd dat hetzelfde vaardigheidsniveau bij alle varianten tot hetzelfde cijfer leidt. Omdat niet alle varianten precies dezelfde moeilijkheidsgraad hebben, kan – bijvoorbeeld – 39 scorepunten bij de ene variant leiden tot het cijfer 6 en bij de andere variant tot het cijfer 7.
Bovendien bewerkstelligt deze systematiek bereikt dat de normering van de rekentoets 2F gelijkwaardig is aan die van de centrale examens rekenen 2F in het MBO. Idem 3F. Voor het VO-leerlingen en MBO-studenten gelden daardoor dezelfde prestatie-eisen.

Een toelichting op deze normeringssystematiek staat op www.cve.nl: kies 'Onderwerpen', vervolgens 'Centrale rekentoets VO' en klik daarna op 'Normering met een vaardigheidsschaal'.